(51) International Patent Classification 3 :		(11) International Publication Number:	WO 84/ 02793
G10L 1/00; H04B 1/66	A1	(43) International Publication Date:	19 July 1984 (19.07.84)

(21) International Application Number: PCT/US83/02030

(22) International Filing Date: 23 December 1983 (23.12.83)

(31) Priority Application Number: 455,250

(32) Priority Date: 3 January 1983 (03.01.83)

(33) Priority Country: US

(71)(72) Applicants and Inventors: HENRICKSON, Larry, Keith [US/US]; 2322 Baker Street, Ames, IA 50011 (US). HUNTINGTON, Dorothy, Allison [US/US]; 830 Escondido Road, Stanford, CA 94305 (US).

(74) Agents: WOODWARD, Henry, K. et al.; Flehr, Hohbach, Test, Albritton & Herbert, Four Embarcadero Center, Suite 3400, San Francisco, CA 94111-4187 (US).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), DK, FI, FR (European patent), GB (European patent), JP, LU (European patent), NL (European patent), NO, SE (European patent).

Published
With international search report.

(54) Title: METHOD AND MEANS FOR PROCESSING SPEECH
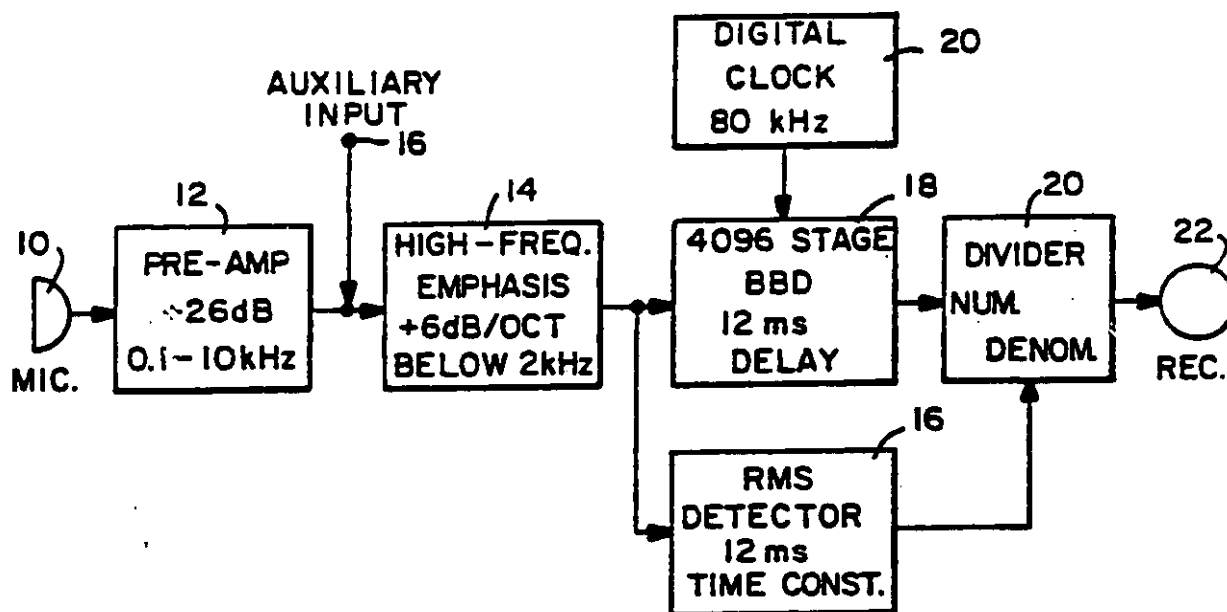
(57) Abstract

A method of and apparatus for processing audio signals in which a measure of amplitude of audio signals in a selected time period is obtained. The audio signals (Fig. 1) for the selected time period are delayed (18) until the measure of amplitude (16) is obtained, and then the delayed audio signals are normalized (20) using the measure of amplitude. High frequency emphasis (14) may be employed prior to obtaining the measure of amplitude. Alternatively, a multi-channel system (Fig. 3) can be employed for processing audio signals in limited frequency bands (32, 34, 36). The method and apparatus are applicable in a variety of applications including hearing aids, audio storage media, broadcast and public address systems, and voice communications such as telephone systems.

-1-

METHOD AND MEANS FOR PROCESSING SPEECH

This invention relates generally to speech processing,
and more particularly the invention relates to a
method and means for amplifying speech, such as for
the hard of hearing, without adversely affecting the
5 signal intelligence thereof.

It is well recognized that persons having sensorineural
hearing impairment generally have a very limited
dynamic range, that is, very little difference between
the intensity level of the softest speech of which
10 they are aware (e.g. speech awareness threshold or
SAT), the intensity level of speech which is most
comfortable for them (most comfortable level or MCL),
and the intensity at which speech becomes too loud to
be tolerated (uncomfortable loudness level or UCL).
15 It is generally agreed that it would be highly desirable
to reduce the wide range of speech intensity levels to
a more restricted range suitable for the sensorineural
hearing impairment of each individual listener.

Speech compression systems are known which employ
20 automatic gain control. However, prior art systems
employing peak clipping and instantaneous compression
produce harmonic distortion which tends to emphasize
the stronger, low-frequency components of speech and
obscures the higher frequencies. A comprehensive

survey is presented by Braida et al in "Hearing
Aids - A Review of Past Research on Linear Amplifica-
tion, Amplitude Compression, and Frequency Lowering",
American Speech-Language-Hearing Association, Rockville,
5 Maryland, April 1979. This survey provides an extensive
critical review of the compression literature in
conjunction with a tutorial on compression concepts.
The survey suggests that the lack of benefits from
compression as shown in the survey literature reflects
10 more a failure of researchers to adequately grasp the
concepts and complexity of compression, in theory and
implementation, rather than the potential benefit of
amplitude compression itself.

It is recognized that the acoustical patterns of
15 speech can be systemically analyzed in three primary
time-domain components: (1) a fine-temporal pattern
reflecting the spectral distribution of each brief
acoustic segment, (2) a gross-temporal pattern reflect-
ing the durations of the various acoustic segments
20 based on changes in fine-temporal patterns, and (3) a
time-varying amplitude pattern. The fine temporal
cues from segments of speech as short as five or ten
milliseconds will often provide a listener with suffi-
cient information to identify the place of articulation
25 for consonants. Similarly, the gross temporal pattern
will often provide sufficient information regarding
the manner of articulation, especially among the
classes of fricatives, affricates and stop-plosives.
The time varying amplitude pattern, or "speech envelope",
30 is the natural result of a speech production process
but may convey mostly redundant information already
conveyed by a gross-temporal pattern. Robinson and
Huntington, in a talk before the Acoustical Society of
America in April, 1973, recognized that conventional
35 compression amplification introduces undesirable
distortion when brief time constants are utilized, and
reacts too sluggishly for longer time constants. A

method was ●oposed in which the aver● power of the
speech wave form over intervals of several tens of
milliseconds is measured continuously and is used to
determine the gain to be applied to the waveform at
5 the center of each interval, with the resulting ampli-
tude compressed signal being delayed by one-half the
length of the averaging interval. Preliminary results
from a computer simulation suggested that speech
intelligibility could be improved by this process.
10 However, further work was not undertaken by Robinson
and Huntington to develop the process.

An object of the present invention is an improved
method of processing speech to facilitate reception
without distorting the intelligible content thereof.

15 Another object of the invention is apparatus for
compressing speech patterns whereby the variations in
time varying amplitude pattern or envelope are minimized
without adversely affecting the fine-temporal and
gross-temporal patterns of the speech.

20 The present invention is directed to a method and
apparatus for processing speech in which a time-
varying averaged or root-mean-square (RMS) amplitude
pattern is obtained and used to normalize the time
varying amplitude pattern of speech and provide a
25 compressed speech pattern positioned between the
speech awareness threshold and the uncomfortable
loudness level, ideally at the listener's most comfort-
able level. Spectral shaping is employed to emphasize
the high-frequency content. The invention can be
30 implemented in a single channel or multi-channel
system. Suitable microphone means is employed to pick
up a speech pattern, and the speech pattern from the
microphone is preamplified and then processed by a
suitable shaping filter which emphasizes the high
35 frequency content thereof. The root-mean-square of

the amplitude of the spectrally shaped signal is then
determined over a specific time period, and the inverse
of the root-mean-square is then used to modulate the
spectrally shaped signal, thus producing a normalized

5   amplitude. Importantly, the shaped signal is delayed
for a sufficient time period to compensate for the
time delay involved in the root-mean-square determina-
tion prior to the amplitude compression. The resulting
signal is thus compressed and then adjusted to the

10   desired hearing range with the spectral shaping provid-
ing a retention of the fine-temporal pattern and the
gross-temporal pattern.

The invention and objects and features thereof will be
more readily apparent from the following detailed

15   description and appended claims when taken with the
drawing, in which:

Figure 1 is a functional block diagram of a single
channel speech processing apparatus in accordance with
one embodiment of the present invention.

20   Figure 2 is a graph illustrating the compression of
speech in accordance with the present invention.

Figure 3 is a functional block diagram of a multi-
channel embodiment of speech processing apparatus in
accordance with the invention.

25   Figures 4A and 4B are functional block diagrams of a
tape recording system in accordance with other embodi-
ments of the invention.

Referring now to the drawings, Figure 1 is a functional
block diagram of a single channel speech processor in

30   accordance with one embodiment of the invention which
has been built using conventional, commercially avail-
able components. In this embodiment a microphone 10

having a bad frequency response (e.an electret
microphone having a response of 100 ‘Hertz to 10K
Hertz such as a Knowles EA 1934) picks up audio signals
and transmits electrical signals to a pre-amplifier 12

5   having 26 dB of gain between 100 Hertz to 10K Hertz.
The amplified signal is then passed to high frequency
emphasis circuity 14 (e.g. TI064 quad amplifier) which
provides 6 dB/octave gain over the range from 100
Hertz to two kiloHertz and a flat response above two

10 kiloHertz. An auxiliary input is provided at 16
whereby signals from a radio receiver, for example,
can be applied to the high frequency emphasis circuitry
14.

The signal from circuitry 14 is then passed to an RMS

15 detector high-frequency emphasis circuitry 14 is also
provided to delay circuitry 18 having a delay equal to
the time constant of the RMS detector 16. In one
embodiment the RMS detector comprised an analog series
AD 536A and the delay circuitry 18 comprised a Reticon

20 SAD 4096 bucket brigade device operated from a 80
kilohertz digital clock 20.

The delayed signal from the delay device 18 is then
applied as the numerator in a divider circuit 20 (e.g.
Analog Devices AD 535 precision divider) and the RMS

25 amplitude of the delayed signal is applied to the
divider 20 as a denominator. Accordingly, the output
from the divider 20 is a delayed amplitude compressed
signal which is applied to the receiver 22 (Knowles ED
1925).

30 Figure 2 is a plot of the compressed output level in
dB SPL for the signal applied to receiver 22 versus
the input level in dB SPL of the signal from the
microphone 10. For input levels below about 45 dB,
the output level is attenuated. At an input level of

35 45 dB, the output level is compressed and maintained

uniform at ●proximately 100 dB SPL wn● is the MCL
level. The compression ratio remains at 10:1 or
greater for input levels above 45 dB.

Figure 3 is a multi-channel signal compression system
5   in accordance with another embodiment of the invention
in which signals are filtered and compressed in a
plurality of frequency bands. In this embodiment
signals from the microphone 30 are applied to a low·
band (100-400 hz) filter 32, a middle band (400-1,600
.10 Hz) filter 34, and a high band (1,600-6,400Hz) filter
36. Signals from each of the filters are passed to
amplitude compressor circuitry 38, 40, and 42. Each
of the compressor circuits includes delay circuitry,
RMS detector circuitry, and divider circuitry as
15 illustrated in Figure 1. Because each channel includes
a narrow band of frequencies, the high frequency
emphasis circuitry of Figure 1 is not required. The
compressed signals are then applied to a summing
amplifier 44 with the composite summed signal then
20 applied to the receiver 46.

Figures 4A and 4B are functional block diagrams of
other embodiments of the invention useful with tape
recorders and in which the compressed signal and the
detected RMS value are both recorded in time sequence
25 with a tape recorder. In Figure 4A signals from the
microphone 50 or other audio source are applied to
amplitude compressor 52 which may be a single channel
device as in Figure 1 or a multiple channel device as
in Figure 3. The compressed audio signal is then
30 recorded in an analog channel of the tape recorder 54,
and the detected RMS value is recorded in an FM channel
of the recorder 54. Thereafter, the recorded compressed
audio signal and the recorded RMS value can be applied
to a multiplier 56 from which the original audio
35 signal and the original dynamic range is produced.
The resulting decompressed signal is applied through

frequency ●-emphasis circuit 58 to t● receiver 59.

Figure 4B is a sampled digital recording system similar
to the analog system of Figure 4A. In this embodiment
signals from microphone 60 are applied to the amplitude
5 compressor 62, as in Figure 4A, and then the compressed
audio signal and the RMS value are converted to digital
form by analog to digital circuits 63 and 65. The
digital signals are then stored in digital recorder
64. The recorded signals are converted back to analog
10 signals by digital to analog converter 57 and multi-
plying DAC 66. The decompressed signals from DAC 66
are frequency de-emphasized at 68 and then applied to
the receiver 69.

These embodiments of the invention are particularly
15 advantageous since tape recorders typically have a
limited dynamic range. Thus, by recording the compressed
audio signal and the RMS on the recorder, the full
dynamic range of the recorded signal can be reconstructed
in the multiplier 56 and multiplier 66.

20 In the preferred embodiments described herein, an RMS
detector has been employed. However, other measures
of the signal amplitude over a period of time, including
an average value and an approximation of the RMS
value, can be employed. As used herein, RMS value
25 includes suitable approximations thereof. Further,
while a divider has been employed in the preferred
embodiments for obtaining the compressed signal, a
logarithmic measure of the detected RMS or averaged
value can be employed for obtaining the compressed
30 value.

The invention has broad applications including, for
example, hearing aids and audio storage media (as
described herein), sampled digital storage system,
broadcast systems, public address systems, and general

voice communication including telephones. The invention
is especially useful for communication in a noisy
environment and through a noisy communication link
such as in field applications.

5 Thus, while the invention has been described with
reference to specific embodiments, the description is
illustrative of the invention and is not to be construed
as limiting the invention. Various modifications and
applications may occur to those skilled in the art
10 without departing from the true spirit and scope of
the invention as defined by the appended claims.

WHAT IS CLAIMED IS:

1. Apparatus for processing audio signals including
speech comprising means for obtaining a measure of
amplitude of speech signals during a selected period
of time, and means for compressing audio signals

5 during said selected period of time based on said
measure of amplitude.

2. Apparatus as defined by Claim 1 wherein said means
for obtaining a measure of amplitude comprises a root-
mean-square (RMS) detector.

10 3. Apparatus as defined by Claim 2 and further including
means for delaying audio signals for said selected
period of time while said root-mean-square detector
obtains the root-mean-square value of amplitude of
said audio signals for said period of time.

15 4. Apparatus as defined by Claim 1 and including
delay means for delaying audio signals for said selected
period of time while said measure of amplitude of
audio signals during the selected period of time is
obtained.

20 5. Apparatus for processing audio comprising
microphone means for receiving audio and generating
electrical signals in response thereto,
high frequency emphasis means connected to said
microphone means for amplifying said electrical signals,

25 amplitude detection means connected to said high
frequency emphasis means for receiving amplified
signals and obtaining a root mean square (RMS) amplitude
of said amplified signals over a selected period of
time,

30 delay means connected to said high-frequency
emphasis means for receiving and delaying spectrally

shaped electrical signals for said selected period of time, and

signal compression means connected to said delay means and to said amplitude detection means and compres-
5   sing the delayed amplified electrical signals in accordance with said root-mean-square average value.

6. Apparatus as defined by Claim 5 wherein said high frequency emphasis means amplifies said electrical signals by six dB per octave at least to 2 kiloHertz.

10   7. Apparatus as defined by Claim 5 wherein said delay means comprises a bucket brigade device.

8. Apparatus as defined by Claim 5 wherein said compression means comprises a divider for dividing said delayed non-linearly amplified signals by said
15   root-mean-square value.

9. Apparatus as defined by Claim 5 and further includ-ing pre-amplification means interconnecting said microphone means to said high-frequency emphasis means.

20   10. In audio signal processing apparatus, a method of compressing the amplitude of signals comprising the steps of

obtaining a measure of amplitude of said signals over a selected time period, and
25   compressing said signals corresponding to said selected period of time in accordance with said measure of amplitude.

11. The method as defined by Claim 10 wherein said measure of amplitude is a root-mean-square value.

30   12. The method as defined by Claim 11 wherein said step of compressing comprises dividing said signals for a selected period of time by the root-mean-square

amplitude ⬤ue thereof.

13.   The method as defined by Claim 10 and further
including the step of emphasizing high frequency comp-
onents of said signals prior to obtaining said root-
5   mean-square value.

14.   Apparatus for processing audio signals comprising
      a plurality of band pass filters for receiving
and filtering audio signals into a plurality of limited
frequency bands,
10   a plurality of amplitude compressor means each
connected with a band pass filter, each of said amplitude
compressor means including
           means for obtaining a measure of amplitude
      of audio signals during a selected period of
15   time,
           means for delaying audio signals for said
      selected period of time, and
           means for compressing the delayed audio
      signals during said period of time based on said
20   measure of amplitude, and
      summing means connected to said plurality of
amplitude compressor means for receiving and summing
compression of audio signals.

15.   Apparatus as defined by Claim 14 wherein said
25 means for obtaining a measure of amplitude comprises a
root-mean-square (RMS) detector.

16.   A method of recording audio signals and obtaining
original dynamic range from the recorded signals, said
method comprising the steps of
30   obtaining a measure of amplitude of audio signals
over a selected time period,
           compressing audio signals corresponding to said
selected time period in accordance with said measure

of amplitude,

    recording the compressed audio signals, and

    recording said measure of amplitude of audio signals.

5   17.   The method as defined by Claim 16 wherein said measure of amplitude is a root-mean-square value.

18.   The method as defined by Claim 16 and further including the steps of

    retrieving the recorded compressed audio signals

10 and the recorded measure of audio signals, and

    decompressing the recorded compressed audio signals using the recorded measure of amplitude of audio signals.

19.   The method as defined by Claim 18 wherein said

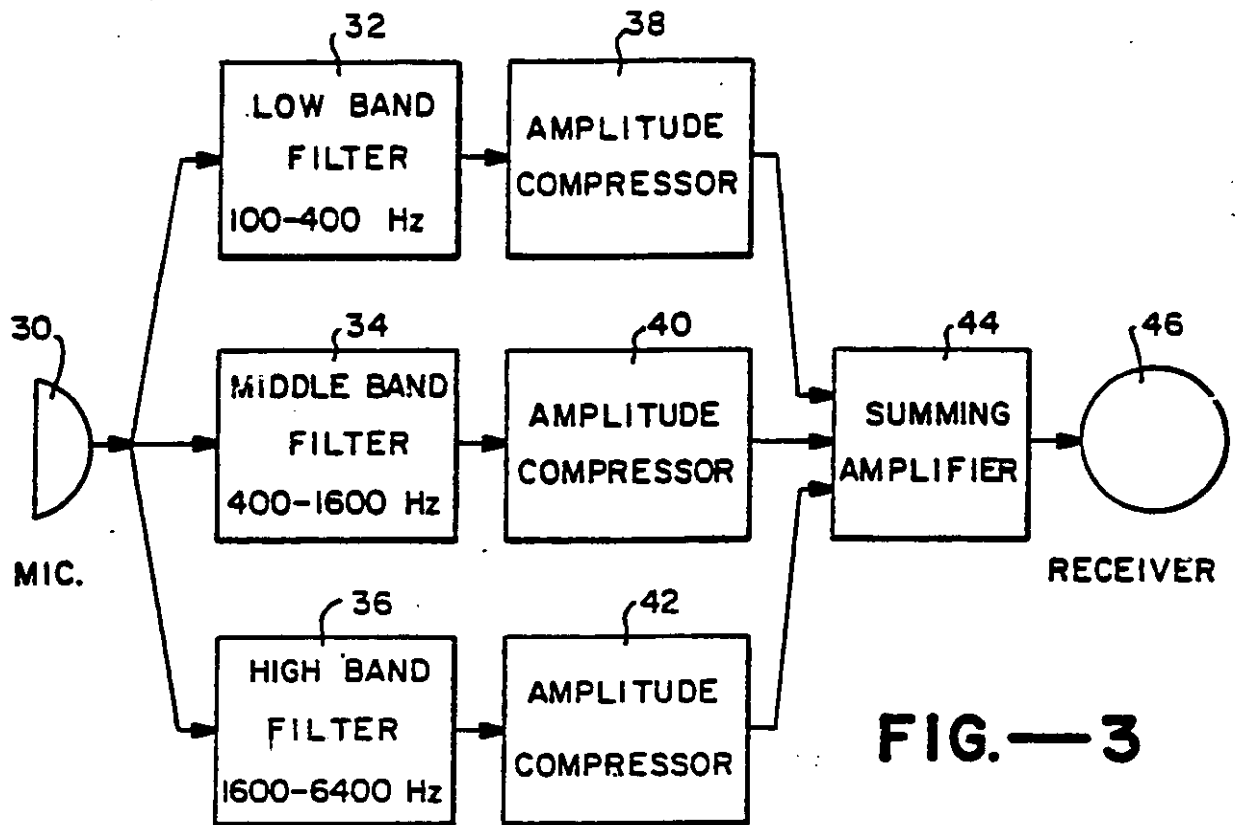15 step of recording includes converting compressed analog signals and measure of amplitude to digital form and said step of decompressing includes converting recorded digital signals to analog form.

20.   The method as defined by Claim 16 wherein said

20 step of recording includes converting compressed analog signals and measure of amplitude to digital form.
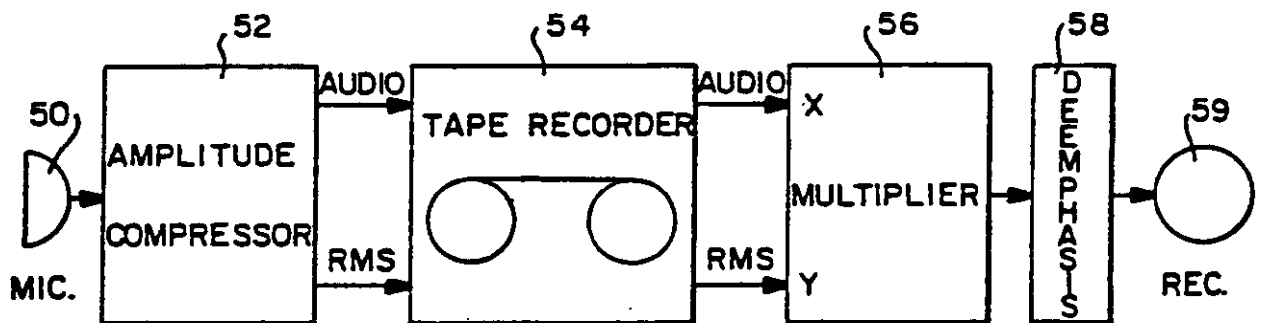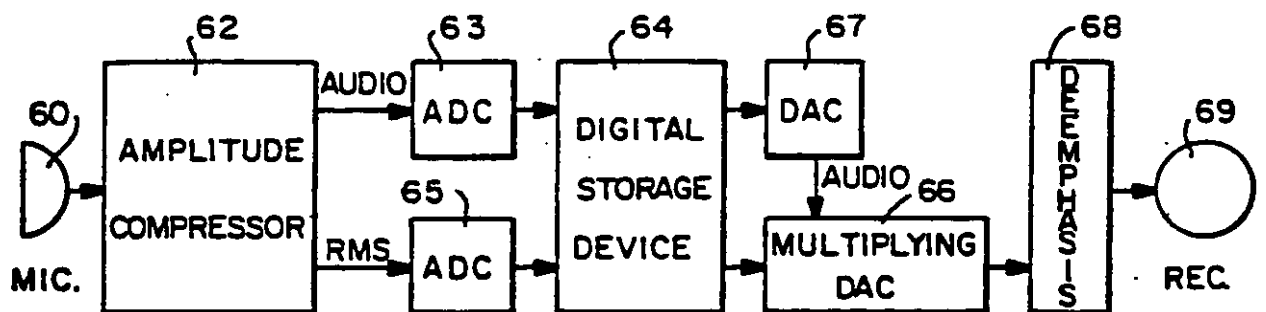
FIG.—3

FIG.—4A

FIG.—4B

SUBSTITUTE SHEET

INTERNATIONAL SEARCH REPORT

International Application No PCT/US83/02030

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) 3

According to International Patent Classification (IPC) or to both National Classification and IPC

US CL. 381-106
INT. CL.3: G10L 1/00; H04B 1/66

II. FIELDS SEARCHED

Minimum Documentation Searched 4

Classification System	Classification Symbols
US	381-106, 110; 179-107.R, 107.FD

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched 5

III. DOCUMENTS CONSIDERED TO BE RELEVANT 14

Category *	Citation of Document, 16 with indication, where appropriate, of the relevant passages 17	Relevant to Claim No. 18
Y	US, A, 4,112,254, 05 September 1978, Blackmer	1-20
Y	US, A, 4,249,042, 03 February 1981, Orban	1-20
Y	Journal of Acoustical Society of America, Vol. 54, No. 1, issued 1978, USA, C.E. Robinson et al, "The Intelligibility of Speech...", see page 314.	1-20

* Special categories of cited documents: 15

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the International filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search 2	Date of Mailing of this International Search Report 2
08 February 1984	15 FEB 1984
International Searching Authority 1	Signature of Authorized Officer 20
ISA/US	E. S. Kemeny

US 0910293909P1

Creation date: 01-04-2005
Indexing Officer: ZREDDA - ZAYED REDDA
Team: OIPEBackFileIndexing
Dossier: 09102939

Legal Date: 11-27-2001

No.	Doccode	Number of pages
1	CTNF	5
2	1449	1

Total number of pages: 6

Remarks:

Order of re-scan issued on ....